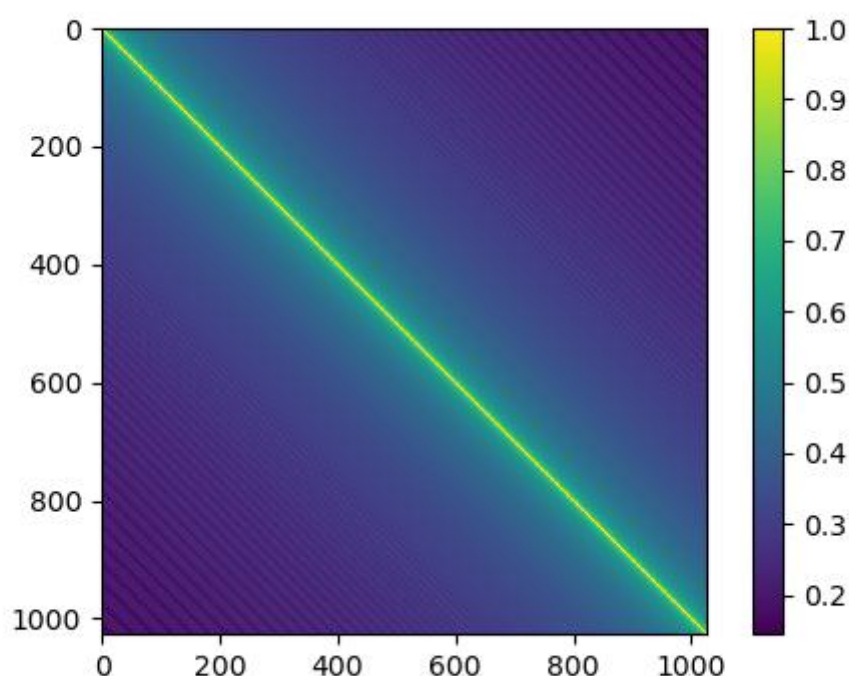


ML HW5 Report

R10922124 林家毅

Problem 1

Visualize the similarity between different pairs of positional embedding and briefly explain the result



上圖是不同 positional embedding pairs 之間的 cosine similarity 所畫出的 heatmap，其中比較亮的區域（接近對角線的部分）是一個 pair 中兩個 positional embeddings 相似度較高的地方，這也代表它們的 positional embeddings 會比較像，而比較遠的則沒那麼像，cosine similarity 較低，透過這樣的關係，可以讓模型利用 positional embedding 知道句子中每個字是在句子的哪個位置。

Problem 2

Clip gradient norm and visualize the changes of gradient norm in different steps. Circle two places with gradient explosion.

